

## 台灣網路使用者檢索行為探析 Exploring Searching Behavior of Network Users in Taiwan

卜小蝶 Hsiao-tieh Pu

世新大學圖書資訊學系副教授

Associate Professor, Department of Library and Information Studies,

Shih Hsin University, Taipei, Taiwan

E-mail: htpu@cc.shu.edu.tw

### 【摘要】

隨著網際網路的蓬勃發展，使用者利用搜尋引擎查詢網路資源已相當普遍，因此藉由分析 Web 搜尋引擎的各種電腦交易記錄 (Transaction Logs)，對瞭解網路使用者的資訊行為應有莫大助益。本文即是以搜尋引擎的檢索詞彙記錄分析 (Search Term Logs Analysis) 為基礎，嘗試觀察台灣地區網路使用者的一些檢索行為特性，包括如檢索詞彙的特性與使用者的檢索需求等，同時探討檢索詞彙分類時的一些問題並提出心得。

### 【Abstract】

With search services over the Web as major sources of information, there is an increasing demand to learn more about what people need and how they search on the net. Through analysis of transaction logs, certain users' search behaviors could be observed. This paper presents the analysis of search term logs from two real-world search engines in Taiwan. Some preliminary results have been obtained, including characteristics of users and their searching needs. Several problems with building taxonomic schemes and classifying search terms are also discussed. The findings are valuable for understanding network information retrieval research, and library & information service applications.

關 鍵 詞：資訊行為；檢索詞彙記錄分析；網路資源檢索

Keywords: Information behavior; Search term logs analysis; Network information retrieval

## 壹、前言

近年來網際網路已成為人們取得資訊的重要管道與來源，使用者藉由如 Yahoo、AltaVista 等類之網路搜尋引擎查詢資訊也相當普遍，因此藉由分析 Web 搜尋引擎的各種交易記錄 (Transaction Logs)，對瞭解網路使用者的資訊行為確實有莫大助益。根據筆者利用搜尋引擎中的檢索詞彙記錄 (Search Term Logs)，分析台灣地區網路使用者的檢索行為研究結果發現 (註 1)，網路使用者有許多特性值得進一步分析研究。本文即是以搜尋引擎的檢索詞彙記錄分析為基礎，嘗試觀察並比較台灣地區與國外網路使用者的一些重要檢索行為特性，包括如檢索詞彙的特性、使用者的特性與檢索需求等，同時並探討檢索詞彙分類時的一些問題並提出心得。有鑑於圖書館及其它資訊服務單位一向關心網路使用者的需求與資源的利用，因此有必要儘早瞭解其檢索需求與特性，以作為提昇資訊服務品質的基礎。此外，隨著中國大陸與各地華文上網人口的快速成長，華文的資訊需求將更為驚人，或許藉由台灣地區使用者的分析，也可作為瞭解華人世界的起步。

圖書資訊學中有關資訊行為的研究 (Information Behavior) 其實已有相當時日，且有豐碩的成果，從早期 ARIST (Annual Review of Information Science and Technology) 一系列對資訊需求與使用 (Information Needs and

Uses) 的探討，陸續有許多重要的理論架構與模式被建構，包括如 Dervin 的意義構成模式 (Sense-Making Model)、Kuhlthau 的資訊查詢過程模式 (Information Search Process Model)、Ellis 的行為模式 (Behavioral Model)、Belkin 的知識狀態變則架構 (Anomalous State of Knowledge Framework)、Taylor 的資訊使用環境 (Information Use Environments)、及 Wilson 的資訊行為模式 (Information Behaviour Model) 等 (註 2)，也有更多的文獻是在探討特定族群的資訊行為，這些研究不僅對瞭解影響資訊行為的因素有所幫助，也奠定了資訊行為研究的基礎。而針對使用者與資訊系統之間互動的檢索行為，也有不少的實証研究與結果產出。(註 3) 然而有關網路使用者資訊行為的研究，目前仍在起步，以下是一些相關研究的介紹。

Catledge & Pitkow 收集使用者端 (client) 瀏覽器 (browser) 所記錄的資料 (client logs)，以了解使用者上網的瀏覽行為，例如使用者所 bookmark 的網址通常不是多數人認定的熱門網站、使用者多數屬於瀏覽者 (browser) 而非檢索者 (searcher) 等 (註 4)；而 Huberman 等人則利用 American Online 所提供的 proxy logs 來統計使用者的瀏覽行為，例如其發現使用者的瀏覽行為部分具有共同的規律，並可利用數學統計方法來預測如某一網站的網頁被造訪的可能次數 (註 5)；Choo 等人則是以質化方法來瞭解網路使用者

的尋求行爲，並結合 F. J. Aguilar 與 D. Ellis 等人的理論，架構網路上的資訊尋求行爲模式（註 6）；而 Spink 等人是以線上問卷調查方式，以瞭解使用者針對一特定檢索主題，是否會進行一連串連續性的檢索行爲（註 7）等。而以檢索詞彙記錄爲分析對象者有 AltaVista 與 Excite 等研究（註 8）（註 9），二者皆是利用大量的檢索詞彙記錄檔案，進行許多有關檢索詞彙長度、檢索詞彙頻率、檢索詞彙相關性（correlation）等統計分析。

上述研究皆是以國外網路使用者爲主，有關華人或台灣地區使用者的研究則相當缺乏。（註 10）筆者曾針對台灣網路使用者的檢索詞彙進行過相關研究，並初步提出一些基本統計與分類整理，本文即是針對檢索詞彙分類過程、與使用者檢索行爲特性提出討論，希望藉以瞭解網際網路環境下使用者的一般檢索行爲與需求特性，以作爲研究網路使用者及網路資訊檢索技術的發展參考（有關研究方法的討論，請參考註 1 所列文章）。以下首先針對檢索詞彙分析中與使用者檢索行爲有關的重要分析結果做一簡介，再對檢索詞彙分類的問題提出討論，最後則是台灣網路使用者特性與檢索需求的探討。

## 貳、檢索詞彙特性分析

筆者曾就台灣地區網路使用者的檢索詞彙記錄進行分析研究，分析的對象主要來自 Dreamer 及 GAIS 所提供

之大量檢索詞彙記錄檔案，Dreamer (D-1998) 的收集期限爲 1998 年約 3 個月的資料，檔案內容記錄包括檢索詞彙及其對應的累計檢索頻率，共有檢索詞彙 228,502 筆，檢索總次數爲 2,183,506 次。GAIS (G-1999) 的收集期限則爲 1999 年約 2 星期的資料，檔案內容記錄則包括檢索詞彙、檢索時戳（timestamp）等。GAIS 的檔案在經過排序整理後共計有檢索詞彙 114,182 筆，檢索總次數 475,564 次。研究項目中與使用者檢索行爲有關的重要結果分別討論如下：

### 一、檢索詞彙簡短

D-1998 的中文檢索詞彙其平均長度爲 6.36 bytes，相當於 3.18 characters，而 G-1999 則爲 7.26 bytes、3.55 characters。顯示多數中文檢索詞彙接近爲一複合名詞（compound nouns）。此外，3.18 characters 約爲雙詞（word bigram or word pair），這與西方所統計的平均長度 2.35 words 相近。（註 11）綜而言之，使用者的檢索詞彙不論東西方皆相當簡短，如此一來，不僅造成使用者的檢索詞彙無法與文件的索引詞彙產生關聯，而順利檢索到相關文獻，同時也難以瞭解使用者的資訊需求或檢索動機，進而設計出滿足其需求的系統。雖然陸續有一些克服簡短檢索詞彙問題的研究，例如利用相似性回饋（Relevance Feedback）或擴大詞彙（Query Expansion）等技術，來增加系統與使用者互動的機會，進而產

生較佳的檢索結果，但目前成效並不明顯。

## 二、少數檢索詞彙佔有多數使用頻率

如圖一所示，在D-1998近23萬筆的檢索詞彙中，檢索頻率超過20次以上者雖僅佔總詞彙數的4.33%，但卻佔檢索總次數的74.89%，而檢索頻率少於（含）3次的詞彙雖佔總詞彙數的78.66%，但其檢索總次數卻僅佔11.86%。顯示台灣地區目前有少數檢索詞彙是使用者熱門的檢索需求，而多數的檢索詞彙則可能是較特殊、冷門、錯誤、或是一些接近熱門詞彙的同義詞。這與AltaVista的研究也有類

似之處，其報告指出檢索頻率低於（含）3者也高達86.4%，顯示低頻詞的比例高是中外皆然；而高頻詞的筆數比例雖然較低，但其檢索總次數卻遠遠超過低頻詞的使用比率，這與所謂的20/80原則頗能呼應，亦即少數熱門詞彙佔有多數的檢索需求。

## 三、以中文為主的檢索詞彙

不論是D-1998或G-1999，中文檢索詞彙都有一定的使用比例（分別為78.20%、78.22%），完全為西文或數字者則佔有不小比例，然而根據抽樣觀察，這些詞彙多數為單詞、且屬於專有名詞，其專指度（Specificity）都

圖一 檢索詞彙頻率分佈

相當高，如“mp3”、“kimo”等，顯示實際上西文資訊的需求並不複雜。然而有趣的是，對一些重要的西文名詞，使用者有時並未使用西文，而是改以中文來輸入，如“Yahoo”以“雅虎”來輸入，顯示使用者仍較習慣以中文來表達其檢索需求。此外，中西文夾雜或其它字元的檢索詞彙比例並不高，且許多詞彙其實也是相關主題的同義詞，如“mp3 下載”其實與“mp3”的需求是類似的。

## 參、檢索詞彙分類探討

為瞭解使用者的檢索需求，將檢索詞彙以其隱含的檢索主題分類（動機分類）是非常重要的嘗試，但首先就牽涉到主題分類的架構設計問題。現行一些網路分類架構或是圖書館界習用的分類架構，並不太適用於檢索詞彙的分類工作。此外，在進行詞彙分類時，由於詞彙過於簡短，如何判斷其動機也是難題之一。以下首先說明檢索詞彙分類架構的設計方式，再針對檢索動機作為分類依據所遭遇的問題提出討論，之後則簡述D-1998近二萬個檢索詞彙的分類統計結果。

### 一、分類目錄架構

現行網路資源的分類方式約可區分為三種。（註12）最常見是依網站內容（Content）來分類，如Yahoo、蕃薯藤等；另一種分類方式是以使用者的喜好作為依據，如PCHome依照各類主題的熱門網站羅列推薦；此外也

有一些網站嘗試以檢索詞彙的角度出發，來設計分類架構，如SearchWords.com就是根據使用者在各大搜尋引擎所鍵入的檢索詞彙記錄，分析後區分為Top Financial、Computer、Sports Sites等網站。這些分類架構其實和圖書館界常用的杜威十進分類法（DDC）的概念相似，只是分析的對象有所不同，一為網路資源，另一為圖書文獻。這種分類架構多半是由上而下（Top-down）的設計方式，也就是預先設定好一套分類體系，再將網路資源或圖書分別歸類到適合的項目。其優點是若在設計時就能考慮到使用者需求，則一般而言，這種架構算是相當簡便、且頗能符合使用者的瀏覽習慣。但其缺點是有些網路資源找不到合適的類別（有書無類），或是類別之間的數量差異很大（有類無書）。再者，使用者對分類架構的認知往往不足，因此即使有多重分類的設計，仍常常迷失在層層的類目之中。此外，由於網路資源的變動極大，這種分類架構更新不易，因此也較缺乏彈性。

此外，網路上的分類目錄系統多是以資源整理者的角度出發，配合現有資源，來揣測網路資源與使用者需求的關係。但根據諸多研究指出，使用者常常不了解其本身的需求，也不了解如何表達其需求，整理者所提供的分類架構與使用者所需求的分類架構往往有所差異。因此，藉由使用者的檢索詞彙的分析，有助於整理者對

使用者需求的掌握，對於設計資源的分類架構與進行分類時也較有幫助。綜合上述，考慮有限人力及時限，本研究首先檢視 D-1998 前 5,000 筆檢索詞彙的主題內容，並收集分析一些具代表性的網路搜尋服務的目錄分類架構，如 Yahoo、Yam、Kimo、Todo、Openfind 等，經討論後暫定一分類架構，再逐筆分類。在第一次整理分類完成後，重新檢討及確定分類架構，並歸納出以下 14 大類及 100 個子類，之後再進行第二次逐筆分類，以確保每個檢索詞彙能經過二次以上的分類檢驗，除了確保較高的正確性與一致性外，也適度檢測分類架構的涵蓋範圍是否足夠。

- 電腦與網路：與電腦軟硬體、網路服務有關的查詢語彙。
- 情與色：與色情、寫真集、或同性戀有關的查詢語彙。
- 影視娛樂：與電影、電視、音樂等娛樂事業有關的查詢語彙。
- 育樂交友：與體育、玩樂、交友聊天有關的查詢語彙。
- 生活消費：與個人生活消費有關的查詢語彙。
- 財經工商：與各種工商產業有關的查詢語彙。
- 旅遊休閒：與旅遊、天氣、交通有關的查詢語彙。
- 教育學習：與學校、考試、補習有關的查詢語彙。
- 遊戲：與電腦軟硬體、網路服務有關的查詢語彙。

- 政治與社會：與政府、社會問題有關的查詢語彙。
- 媒體與出版：與各類傳播媒體、出版事業有關的查詢語彙。
- 醫療保健：與醫院、疾病健康有關的查詢語彙。
- 人文藝術：與文學、歷史、藝術有關的查詢語彙。
- 學術資訊：與科學技術、圖書資訊有關的查詢語彙
- 無法區分者為其它

上述架構其實是以由下往上 (Bottom-up) 的方式來設計，由於使用者的檢索詞彙內容多元新穎 (註 13)，且主要的需求範疇與網路資源所提供的內容並不一定相符，因此若要延用或參考現行各種分類架構並不適用，最好是由使用者的檢索詞彙中，分析歸納整理出合適的分類範疇。因此本研究所提供的分類架構與圖書館中的分類系統或是現行的網路分類目錄架構有所不同。

## 二、以檢索動機為分類依據

在進行詞彙的分類時，主要是以使用者動機為考量，而非詞彙本身的主題性質，例如“mp3”主要歸類在影視娛樂類，而非電腦&網路類。因為多數查詢“mp3”的使用者，其主要目的是在下載各種mp3的歌曲，想研究軟體技術者較少。但為避免一詞多義所造成的可能誤解情形，也為了讓分類能較為彈性以符合多元的檢索需求，多重分類有其必要性，包括分

類架構中的類目參照，與檢索詞彙的多重分類。例如電腦&網路類中的BBS子類與育樂交友的聊天交友子類、或是教育學習的學校子類等就有相當的關聯性，而“mp3”這個檢索詞彙也有必要作參照分類至電腦&網路類。但即使提供多重分類，仍有一些問題待克服，包括詞彙簡短難以確認其動機與目的，或是不同使用者對同一檢索詞彙的認知層次有所不同而應有所區別（如何過濾）等。對於動機の確認方式，雖然在分類時有一定的主觀判斷，但每一詞彙都需經過兩人檢視同意，且有疑義時，必須上網利用搜尋引擎查詢，以使用者所能獲得的檢索結果來判斷其可能的用途，進而推測其可能的動機與目的。

此外，分類的目的主要是將較大的集合區分成較小的集合，以方便使用者瀏覽及整理者管理，但網路使用者的需求通常是屬於主題性質居多，只提供分類並不能滿足其需求。以所分析的近2萬筆檢索詞彙中，有超過六成是屬於主題性的檢索詞彙，雖然有近四成是查詢單位、網站、或人名等專有名詞，但其中有許多詞彙是屬於主題類型的查詢，例如“鐵路局”與“火車票”，其實其檢索需求是相當接近的。若要真正滿足使用者的檢索需求，主題的分析是相當重要的，而且主題的分析也不能只考慮單面向（facet），多重面向的分析有其必要性。例如以下是一些有關旅遊類的檢索詞彙，約略可以歸納出地圖、旅

館、景點、機票、及特殊活動如釣魚等幾個主題，而這些主題之間存在許多語意及動機上的關聯，例如若要滿足“日本”這個檢索詞彙的需求，上述幾個主題都應納入考量。

· 地圖、日本、墾丁、飯店、航空公司、航空、溫泉、機票、澎湖、香港、新竹、eva、宜蘭、長榮航空、長榮、福華飯店、花蓮、電子地圖、救國團、台中、新加坡、阿里山、台南、釣魚、香格里拉、福華、旅館、加拿大、高雄、國內旅遊、遠東航空、古蹟、凱撒、綠島、風景、觀光局、淡水、華航、台北市、台灣、中華航空、泰國、復興航空、旅行、自助旅行、澳洲、桃園、陽明山、飛機等。

然而要從如此龐雜的檢索詞彙進行多面向分析，的確相當困難，但根據上述統計，其實少數的檢索詞彙佔有多數的使用頻率，若能先針對這些高頻詞為範疇，將每一類的檢索詞彙分析歸納出其面向特徵，之後再進行主題的分析，應是可行方式。例如上述旅遊類就可能包括地圖、地點、活動等面向特徵，若配合主題的分類，應能滿足多數有關旅遊類的檢索需求。

### 三、主題分類統計

就檢索詞彙的分類結果如圖二所示，若依檢索次數統計，查詢電腦&網路類者最多，佔近20%，超過情&色等熱門類別；而查詢學術資訊類最少，僅約佔1.4%。若依檢索筆數統

## 圖二 檢索詞彙大類統計（依檢索次數排名）

計，電腦&網路類的檢索詞彙筆數仍佔最多（15.84%），而學術資訊類的筆數 2.84%仍敬陪末座，其餘大類的排行與按次數排列雖略有差異，但並不明顯。

而就近一百類的子類統計來看，每類檢索詞彙除了可以觀察其多元化的程度外，也可以了解其中重要的主題需求為何，但限於篇幅，在此僅略舉數例，詳細統計可參考附件一。例如人文藝術類中以檢索"小說"比例最高，其餘的主題並不明顯，而人名的檢索詞彙中以"金庸"比例最高，查詢現代作家的比例並不高。雖然人文藝術類中有關人名的查詢比例相較其它許多大類高，但整體而言其檢索詞彙

並不多元。又如電腦&網路類中以查詢各類下載軟體最多，這些軟體多半為個人用途的工具型軟體，如輸入法、繪圖、防毒軟體等，較具獨特性的軟體並不多見，這對發展軟體的產業而言似乎不算是太好的消息。此外，電腦&網路類的英文檢索詞彙是所有類別中佔有比例最高者，這與電腦族群使用原文術語習慣一致。另外，其中的網路服務子類中多數是查詢免費 E-mail，這說明 E-mail 服務確實是相當成功的網路服務項目之一。至於BBS子類中則以查詢BBS網址為主。雖然電腦&網路大類的檢索詞彙相當多，但檢視詞彙內容，其需求並不複雜。



## 肆、台灣網路使用者特性與檢索需求分析

依據對檢索詞彙的主題內容分析，及比較國內外一些相關研究，以下是有關台灣網路使用者的一些重要特性與檢索需求探討。

### 一、年輕化的使用族群

由檢索詞彙的內容分析中，有許多詞彙屬於時下的流行用語，深入專指的主題並不多見，即使有一些專業資訊檢索需求出現，也多偏向課程、考古題、論文等課業性需求，這些多少透露出使用者中有相當大的比例為年輕族群。這與蕃薯藤或是資策會的問卷調查報告結果，網路族平均年齡集中在 18-29 為主，頗能呼應。（註 14）但相對於國外調查如 Gvu Survey（註 15），平均年齡為 35 歲則相差甚大。所幸這種現象應會隨著上網人口與需求成熟度的增加而改善，例如蕃薯藤 1998 年調查平均年齡為 24.5，到了 1999 年則升至 25.3，一些 D-1998 的檢索詞彙在一年後的 G-1999 中排名上昇，也出現一些較專指的主題詞彙。

### 二、共同興趣，中外皆然；流行事物，中外不盡相同

由 D-1998 及 AltaVista 的前 25 大檢索詞彙可發現，中外共同熱門的檢索主題包括色情 (sex)、搜尋引擎 (search engine)、mp3、圖片 (photo)、及遊戲 (game) 等。但隨風土民情不同，

則有不同區域性的檢索主題，例如國外找賀卡 (greeting card)、買禮物 (gift) 的比例相當高，但在台灣則較不流行。而台灣相當流行的寫真集、日本文化也是西方較不易見到。

### 三、不同區域的檢索詞彙使用情形仍有相當大的差距

此外，值得注意的是，AltaVista 的前 25 大檢索詞彙佔總檢索次數的 1.5%，而 D-1998 則佔了 14.4%，顯示不同區域的檢索詞彙其使用情形仍有相當大的差距。而影響檢索詞彙的使用因素很多，至少包括網路資源的規模、東西方語文特性的不同、及網路使用人數的多寡等。就網路資源的規模而言，據估計，目前全球約有 21 億個 Web 網頁（註 16），而據瞭解，台灣的網頁目前約只有 400 萬頁，比例不到國外資源的千分之 2（註 17）；此外，從 AltaVista 的 distinct query 約有 1 億 5 千萬筆，排名最高的檢索詞彙也佔不到 0.05%，而中文的檢索詞彙相形之下並無如此之多樣性，且熱門詞彙佔有相當大的比例，或許是因為網路內容的不夠充足，而造成使用者查詢意願相對降低。除了網路資源的不足，語言特性不同也可能影響檢索詞彙的多樣性，舉例來說，相較於英文，中文一般而言其詞彙的語意較為明確，且無大小寫、詞類等變化，因此用以表達同一概念的檢索詞彙的變異性相對可能較少。最後，使用者數量的多寡自然也會讓檢索詞彙的使用

造成變化，同時目前使用者未將網路視爲主要資訊尋求管道也是影響使用情形差異的重要因素之一。

#### 四、言行有別

根據蕃薯藤的問卷調查報告指出，網路族最常去瀏覽的網站類型中，情色網站僅佔 0.8%、學術網站佔有 3.8%，和本研究檢索詞彙分析中情 & 色、學術資訊類的 16.9% 與 1.4% 有些差異。其實在國外也有類似的情形，例如同樣是 Excite 的使用者，一爲根據線上調查 (Survey)，另一種方式是利用檢索詞彙分析 (Search Term Logs)，其檢索行爲也有言行不一的情形產生。(註 18) 一般而言，調查方法 (Survey) 不論是以問卷、面談、或是觀察方式來進行，都有較爲主觀與取樣有限等限制，但較能深入瞭解與解釋個人行爲；而電腦記錄分析 (Transaction Log Analysis) 則具有較客觀、採樣規模較大等優點，但也較不容易探究使用者的真實動機與行爲。若欲瞭解使用者檢索需求的全貌，這兩種方式有必要整合。

#### 五、學術性資訊需求低

由上述說明，台灣網路使用者中以年輕使用者居多，且其它語文的檢索詞彙比例相當低。在近 2 萬筆的檢索詞彙中，學術資訊類的檢索次數敬陪末座且比例很低，且檢視其詞彙內容，多屬於廣泛概念性、專指性不足的 non-distinguishable 詞彙，例如科學

技術類，扣除單位之類的專有名詞，前十大主題性詞彙爲 " ufo "、" 數學 "、" 天文 "、" 海洋 "、" 地理 "、" 衛星 "、" 科技 "、" 物理 "、" 化學 "、" 機械 " 等，又如與圖書館資料查詢有關的檢索詞彙如 " 論文 "、" 文章 "、" 圖書 "、" 期刊 "、" 研究報告 " 等。而且即使是中頻或是低頻的檢索詞彙，其多樣性也不明顯。或許限於台灣網路資源量的不足與高品質資訊有限，許多具有學術性需求的使用者目前仍未視網路爲主要途徑，仍透過圖書館或資料中心之類的單位來查詢國內外資料。這種情形是否會隨著網路資源規模的成長而有所改善，值得觀察。

#### 伍、結語

隨著網際網路逐漸成爲人們取得資訊的重要管道與來源，有關網路使用者的資訊行爲研究也日受重視。而圖書館及其它資訊服務單位向來關心網路使用者的需求與資源的利用，因此有必要儘早瞭解其檢索需求與特性，以作爲提昇資訊服務品質的基礎。本文即是藉由分析搜尋引擎的檢索詞彙記錄，初步獲得一些有關台灣地區網路使用者的檢索行爲特性與需求，並探討檢索詞彙分類的問題與心得。綜而言之，網路使用者的檢索詞彙一般而言相當簡短，且少數檢索詞彙佔有多數使用的頻率。此外，現行圖書館或網路上的分類架構並不適用於檢索詞彙的分類工作，如何設計出

以檢索動機作為分類依據的架構則需要進一步研究。而就網路使用者的特性及需求而言，相較於國外使用者與檢索需求的多元化，台灣網路使用者具有年輕化的特性，學術性需求不高，且整體的檢索行為也反映出資源與人數規模的不足，亟待建設與經營。

## 陸、誌謝

感謝 Dreamer 及 Openfind 所提供的大量檢索詞彙記錄檔案，及世新大學圖書館高雅慧小姐，圖書資訊學系陳文生、賴國峰、許賢仁、楊翔純、及蕭志遠等五位同學的辛苦分類。此外，中研院資訊所 Csmart 小組莊水龍先生的技術支援也使得本研究能順利進行，在此一併誌謝。

## 註釋

- 註 1：卜小蝶，「網路使用者查詢語彙記錄之分析與應用初探」，資訊傳播與圖書館學 6 卷 2 期 (民國 88 年 12 月)，頁 49-62。
- 註 2：B. Dervin, "An Overview of Sense-Making Research: Concepts, Methods and Results to Date," International Communication Association Annual Meeting, 1983.  
 <<http://communication.sbs.ohio-state.edu/sense-making/art/artdervin83.html>>; C. C. Kuhlthau, "Inside the Search Process: Information Seeking from the User's Perspective," Journal of the American Society for Information Science 42:5(1991): 361-371; D. Ellis, "A Behavioural Approach to Information Retrieval Systems Design," Journal of Information Science 15 (1989): 237-247; N.J. Belkin et al., "ASK for Information Retrieval: Part I. Background and History," Journal of Documentation 38:2(1982): 61-71; R.S. Taylor, "Information Use Environments," Progress in Communication Sciences 10(1991): 217-255; T. Wilson and C. Walsh, "Information Behaviour: An Inter-Disciplinary Perspective," 1996. <<http://www.shef.ac.uk/uni/academic/I-M/is/publications/prelims.html>>.
- 註 3：可參考 A. Coleman, "Information Seeking Behavior - A Bibliography," Feb. 6 2000 updated. <<http://www.rscdd.org/home/acoeman/inforseekbeh.html>>.
- 註 4：L. D. Catledge and J. E. Pitkow, "Characterizing Browsing Strategies in the World Wide Web," WWW'95, 1995. <<http://www.igd.fhg.de/www/www95/papers/80/userpatterns/UserPatterns.Paper4.formatted.htm>>.
- 註 5：B. A. Huberman et al., "Strong Regularities in World Wide Web Surfing," Science 280:5360 (1998): 94-97.
- 註 6：Chun Wei Choo, Brian Detlor, and Don Turnbull, "Working The Web: An Empirical Model of Web Use," 33rd Hawaii International Conference on System Science (HICSS), 2000.
- 註 7：Amanda Spink et al., "Searching the Web: a Survey of EXCITE Users," Internet Research:

Electronic Networking Applications and Policy 9(1999): 117-128.

- 註 8：C.Silverstein et al., "Analysis of a Very Large AltaVista Query Log," SRC Technical Note, 1998. <<http://www.research.digital.com/SRC/>>.
- 註 9：B. Jansen, Amanda Spink, and Telfko Saracevic, "Real Life, Real Users, and Real Needs: a Study and Analysis of User Queries on the Web," Information Processing & Management 36 (2000): 207-227.
- 註 10：林珊如、洪曉珊，「網路書店使用者資訊行為之研究」，資訊傳播與圖書館學 5 卷 4 期（民國 88 年 6 月），頁 21-33；吳美美，「中文資訊檢索系統的使用和使用者的檢索問題探討」，資訊傳播與圖書館學 6 卷 2 期（民國 88 年 12 月），頁 23-38；黃慕萱，「停頓行為探討--以國立臺灣大學之終端使用者為例」，圖書館學刊 13 期（民國 87 年 12 月），頁 61-86；陳佳君，「檢索詞彙來源探析」，大學圖書館 2 卷 3 期（民國 87 年 7 月），頁 27-57；李宜容，「人文及社會學科讀者使用線上公用目錄檢索詞彙之研究」，大學圖書館 2 卷 3 期（民國 87 年 7 月），頁 72-104。
- 註 11：同註 9。
- 註 12：DESIRE: Development of a European Service for Information on Research and Education, 1997. <[http://www.lub.lu.se/desire/radar/reports/D3.2.3/class\\_v10.html](http://www.lub.lu.se/desire/radar/reports/D3.2.3/class_v10.html)>；陳樹年，「搜索引擎及聯絡信息資源的分類組織」，圖書情報工作 2000 年第 4 期，頁 31-37。
- 註 13：根據 D-1998 的前近 2 萬個檢索詞彙與國語會所發行超過二十萬詞的國語詞典作一比較，符合比例僅接近二成。
- 註 14：蕃薯藤調查網，99 台灣網路使用調查結果初步統計，<<http://survey.yam.com.tw/99/99result.htm>>；資策會網際網路情報服務中心，88 年我國網際網路應用調查，<<http://www.find.org.tw/survey/survey1999.asp>>。
- 註 15：GVU's WWW User Surveys. <[http://www.gvu.gatech.edu/user\\_surveys/](http://www.gvu.gatech.edu/user_surveys/)>.
- 註 16：Cyveillance, Inc., "Sizing the Internet," July 10, 2000. <<http://www.cyveillance.com/newsroom/3012.asp>>.
- 註 17：根據另一項由 Vilaweb (7/2000) 的統計調查報告指出，全球網頁以英語佔有近 7 成比例 (68.39%) 高居第一，且遙遙領先其它語系，而中文網頁則佔有 3.87% (超過 1,200 萬頁)，排名第四。<<http://www.vilaweb.com/>>。
- 註 18：註 7 與註 9 分別以線上調查 (Survey) 與檢索詞彙分析 (Search Term Logs) 兩種不同方式來瞭解 Excite 使用者的檢索行為，有關 sex 的檢索詞彙在後者佔有一定比例，但卻未出現在前者的統計結果中。





